Visual Analytics para la Detección Temprana de Intrusiones en Tiempo Real

Mariños Hilario, Princce Yorwin

Orientador: Prof Dr./Mag./Ing. Nombre del Asesor

Plan de Tesis presentado la Escuela Profesional Ciencia de la Computación como paso previo a la elaboración de la Tesis Profesional.

UNSA - Universidad Nacional de San Agustín de Arequipa
Junio de 2025

Índice

1. Introduccion

A pesar de los avances en los sistemas de detección de intrusiones (IDS) mediante técnicas de aprendizaje automático, la literatura reciente señala una limitación crucial: la falta de adaptabilidad de estos sistemas ante patrones de tráfico malicioso emergentes en entornos dinámicos y en tiempo real [1, 2]. Los métodos tradicionales operan sobre conjuntos de datos estáticos o procesan lotes de información previamente recolectada, lo que impide detectar de forma oportuna comportamientos nuevos o no observados durante el entrenamiento inicial. Esta deficiencia representa un riesgo significativo en ciberseguridad, ya que los atacantes suelen modificar sus tácticas para evadir mecanismos automatizados de detección.

Ambos trabajos analizados —*ILIDViz: An Incremental Learning-Based Visual Analysis System for Network Anomaly Detection* [1] y *FSA-IDS: A Flow-based Self-Active Intrusion Detection System*— coinciden en señalar que los IDS actuales requieren capacidades de aprendizaje continuo (incremental) y mecanismos que reduzcan la dependencia del etiquetado manual. ILIDViz introduce una solución basada en visualización interactiva y aprendizaje incremental para mejorar la comprensión del modelo por parte de los analistas humanos, mientras que FSA-IDS propone un sistema autoactivo basado en flujos, que minimiza el esfuerzo humano necesario para mantener actualizado el sistema. Sin embargo, ambos estudios también revelan que persiste una brecha crítica en el diseño de herramientas que permitan observar de forma clara y efectiva cómo evolucionan los patrones maliciosos en el tiempo y cuándo el sistema debería adaptarse.

En este contexto, se identifica un problema específico: los sistemas de detección de intrusiones actuales no permiten visualizar de forma efectiva la aparición de nuevos patrones de tráfico malicioso en tiempo real, lo que dificulta la intervención del analista y limita la capacidad del modelo para adaptarse a amenazas emergentes. Esta situación compromete la capacidad de anticipación y respuesta de los Centros de Operaciones de Seguridad (SOC), donde los analistas deben tomar decisiones rápidas con información frecuentemente incompleta o poco interpretable.

Frente a este problema, el presente trabajo plantea como objetivo general diseñar una solución de análisis visual que permita detectar y visualizar en tiempo real la aparición de patrones de tráfico malicioso desconocidos, facilitando la intervención del analista y mejorando la adaptabilidad del sistema de detección. Esta solución buscará combinar técnicas de visualización interactiva con criterios de incertidumbre del modelo, de manera que los expertos puedan identificar rápidamente eventos atípicos o ambiguos que podrían representar nuevas amenazas.

Objetivos específicos

- Analizar las limitaciones de los sistemas IDS actuales en cuanto a su capacidad para adaptarse a datos nuevos en escenarios dinámicos.
- Diseñar una interfaz visual que muestre en tiempo real la evolución del tráfico y

destaque flujos maliciosos emergentes o inciertos.

- Implementar visualizaciones que permitan comparar muestras benignas y sospechosas a lo largo del tiempo, apoyando el diagnóstico humano.
- Evaluar el impacto de la herramienta en la rapidez y precisión del análisis frente a métodos tradicionales basados en métricas numéricas o tablas.

2. Descripción del Dataset

El dataset *CICIDS2017* es un conjunto de datos especializado en el análisis de ciberseguridad, particularmente enfocado en la detección de tráfico malicioso dentro de redes informáticas. Fue diseñado para simular un entorno realista donde se registran sesiones de conexión tanto benignas como maliciosas, lo que permite entrenar y evaluar sistemas de detección de intrusos. Este dataset es ampliamente utilizado en la investigación de amenazas avanzadas, ya que contiene más de 80 características extraídas de flujos de red capturados mediante herramientas como *CICFlowMeter*.

Cada fila del dataset representa un flujo de conexión entre dos extremos (cliente-servidor), incluyendo variables como duración, número de paquetes, tamaños, tasas de envío, estadísticas temporales, indicadores de banderas TCP y, finalmente, una etiqueta que indica si el flujo es benigno o pertenece a un tipo específico de ataque (DoS, PortScan, BruteForce, entre otros). Esto permite realizar análisis supervisados y no supervisados en el ámbito de la detección de amenazas.

Cuadro 1: Resumen de técnicas aplicadas al dataset *CICIDS2017*

Técnica aplicada	**Descripción**	**Variables involucradas**	**Valor que aporta**
Preprocesamiento de datos	Limpieza de valores nulos, conversión de tipos, detección de outliers y duplicados	Todas las variables	Asegura consistencia y evita errores durante el análisis o modelado
Clasificación de tipos de datos	Identificación de variables continuas, discretas y categóricas	Todas las columnas	Permite aplicar métodos estadísticos adecuados a cada tipo de dato
Análisis estadístico	Cálculo de media, desviación estándar, mediana, moda, correlación y covarianza	Variables numéricas	Facilita la identificación de patrones y relaciones entre características
Análisis de balance de clases	Exploración de la variable *Label* para evaluar la distribución de etiquetas	Label	Detecta desequilibrios entre tráfico benigno y malicioso
Tratamiento de valores nulos	Eliminación o imputación de registros con valores nulos en columnas clave	Flow Bytes/s, Packet Length Mean, etc.	Evita distorsión en las métricas y errores en el modelado
Distribución temporal y granularidad	Evaluación de timestamps para identificar variaciones temporales y estacionales	Timestamp, Flow Duration	Permite detectar tendencias y variabilidad en diferentes períodos temporales
Extracción de features relevantes	Selección de columnas más representativas para el análisis	Variables correlacionadas con *Label*	Optimiza el rendimiento y la interpretabilidad en modelos de detección

3. Preguntas

El dataset CICIDS2017 contiene datos de tráfico de red con más de 80 variables para analizar tráfico benigno y ataques (DoS, PortScan, BruteForce, etc.).

Variables clave usadas en el análisis:

- **Flow Duration:** duración del flujo de conexión (importante para medir tiempo activo).
- **Fwd IAT Total y Bwd IAT Total:** tiempos entre paquetes enviados hacia adelante y hacia atrás (clave para distinguir tipos de tráfico).
- **Label:** clase que indica si el flujo es benigno o malicioso.
- **Timestamp:** marca temporal para análisis de tendencias y comportamientos temporales.

Este contexto ayuda a entender cómo se comporta el tráfico y qué variables son críticas para detectar ataques.

4. Problemas identificados en el dataset CICIDS2017

- Desbalance extremo en la variable `Label`: 98 % benigno, menos del 2 % ataques.
- Presencia de valores negativos erróneos en `Flow Duration`, corregidos para evitar distorsión.
- Muchas variables con alta proporción de ceros o datos faltantes (flags TCP, métricas bulk), descartadas por falta de información.

5. Descubrimientos al analizar los datos

- Variables temporales como `Flow Duration`, `Fwd IAT Total` y `Bwd IAT Total` están altamente correlacionadas, mostrando que la duración y tiempos entre paquetes son factores clave.
- Variables con poca o ninguna variabilidad fueron descartadas.
- El análisis estadístico y gráfico confirma diferencias claras entre tráfico benigno y malicioso.

6. Patrones de tendencia detectados

- El análisis de la variable `Flow Duration` a lo largo del tiempo muestra fluctuaciones significativas según la hora del día.
- Esto sugiere la existencia de picos de actividad relacionados con el comportamiento humano o ataques programados.
- El uso del campo `Timestamp` permitió hacer un análisis temporal agregando los datos por hora.

7. Comportamiento humano y movilidad en relación con la geografía

En este dataset, no contamos con datos geográficos explícitos, pero podemos usar atributos como `Protocolo`, `Puertos de destino` y `Horario` para inferir patrones de actividad humana en la red.

Los gráficos muestran que el tráfico varía según protocolo y puerto, y que la actividad maliciosa tiene horarios específicos, reflejando comportamientos planificados o estacionales.

Referencias

[1] X. Tian, Z. Wu, J. Cao, S. Chen, and X. Dong, "Ilidviz: An incremental learning-based visual analysis system for network anomaly detection," *Virtual Reality & Intelligent Hardware*, vol. 5, no. 6, pp. 471–489, 2023.

[2] L. T. Nguyen, J. Kwon, H. Kim, and S. Cho, "Fsa-ids: A flow-based self-active intrusion detection system," *IEEE Access*, vol. 10, pp. 88256–88269, 2022.